ACTIVITY

#1:

- Take a penny. - Flip the penny 40 times. - Record the number of heads and tails in the chart below.

HEADS	TAILS
卌 卌 卌 卌 |	卌 卌 卌 ||||

- Calculate a 95% confidence interval for *p* (the proportion of ~~tails~~ heads flipped)

$\hat{p} = \frac{21}{40}$ $(0.37025, 0.67975)$

- Perform a test to see if your coin is fair
  - *Hint: your coin is claimed to be fair. Thus, the proportion of tails is supposed to be? TEST THIS*

$H_0: p = 0.5$ $z = 0.316$

$H_a: p \neq 0.5$ p-val $= 0.752$

Fail to reject → $p = 0.5$ → coin is fair

ACTIVITY

#1:

- Take a quarter. - Flip the penny 40 times. - Record the number of heads and tails in the chart below.

HEADS	TAILS
~~||||~~ ~~||||~~ ~~||||~~ ~~||||~~ |	~~||||~~ ~~||||~~ ~~||||~~ ||||

- Calculate a 95% confidence interval for *p* (the proportion of ~~tails~~ heads flipped)

$$\hat{p} = \frac{21}{40} = (0.37025, 0.67975)$$

- Perform a test to see if your coin is fair
  - *Hint: your coin is claimed to be fair. Thus, the proportion of tails is supposed to be? TEST THIS*

$H_0: p = 0.5$ $\quad z = 0.3162$

$H_a: p \neq 0.5$ $\quad$ p-val $= 0.752$

* fail to reject $\rightarrow p = 0.5$

Comparing Two ~~Means~~ Prop.

- We want to compare... 2 populations use 2 samples
- Each group is considered... a distinct sample from its pop.
- Responses in each group are... independent

2 Proportion Z Confidence Interval

We are comparing... 2 proportions

If the 2 proportions are the same, then... $p_1 = p_2$ and $p_1 - p_2 = 0$

So we are looking at... difference btw. the 2 proportions

$\frac{16}{40}$ $\frac{16}{40}$

Interval Formula:

GENERIC:

Statistic ± (critical value)(std. dev. of statistic)

FOR 2 PROPORTIONS SPECIFICALLY:

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \underbrace{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

P Q P Q

Parameters: p_1 and p_2 $p_1 - p_2$

Statistics: $\hat{p}_1$ and $\hat{p}_2$ $\hat{p}_1 - \hat{p}_2$

Standard Error = → std. error = std. dev. of stat

INTERPRETATION:

We are ___ % conf. that the difference btw. the prop. of pop. 1 and pop 2. is btw. __a__ and __b__.

$\hat{p}_P = \frac{21}{40}$ $\hat{p}_Q = \frac{21}{40}$ $(-0.2189, 0.2189)$

$1\ \frac{19}{40}$ $2\ \frac{22}{40}$

2 Proportion Z Test

Same steps for the test of significance:

1. Assumptions
2. Hypotheses
3. Test Statistic
4. p-value
5. Conclusion

2 populations with each of their statistics and parameters… (denoted with numbers)

	Pop. 1	Pop. 2
population ~~mean~~ prop.	p_1	p_2
~~population std. dev.~~		
sample size	n_1	n_2
sample ~~mean~~ prop.	$\hat{p}_1$	$\hat{p}_2$
~~sample std. dev.~~		

Hypotheses:

- We are comparing... 2 population proportions

Ho: $p_1 = p_2$

Ha: $p_1 \gtrless p_2$, $\neq$

* no #'s

Ho: $p_1 - p_2 = 0$

OR

Ha: $p_1 - p_2 \gtrless 0$, $\neq$

Ho: $p = 0.2$

Test Statistic:

GENERIC FORMULA:

$$\frac{\text{statistic} - \text{parameter}}{\text{std. dev. of statistic}}$$

$H_0: p_1 = p_2$

FOR 2 PROPORTIONS SPECIFICALLY:

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\underbrace{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}_{\text{Standard Error}}}$$

$$\hat{p} = \frac{X_1 + X_2}{n_1 + n_2} \neq \hat{p}_1 + \hat{p}_2$$

$$\frac{80}{100} + \frac{40}{100} = \frac{120}{200}$$

Standard Error =

Notice how this is diff. from Conf. Int

$$\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Why? What is this called? pooled

b/c $H_0: p_1 = p_2$

What is pooled?

Combining together info

$H_0: p_1 = p_2$

$$\sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}}$$

$$\sqrt{p(1-p)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

$\hat{p}_1 \neq \hat{p}_2$

$\frac{19}{40}$ $\frac{23}{40}$

$\frac{19+23}{80}$

Why do we do this test pooled?

b/c $H_0: p_1 = p_2$

P-Value:

Same as... before

$P(z \gtrless \underline{\text{test statistic}})$

* for $\neq$, 2x

$H_0: p_1 = p_2$

$H_a: p_1 \gtrless_{\neq} p_2$

Conclusion:

Same as... before

- 2 sentences:

- We reject/fail to reject b/c p-value.....

- We have suff. evidence that the prop of $\underline{\quad 1 \quad}$ is $\overset{=}{<} \gtrless_{\neq}$ $\underline{\text{prop of 2}}$.

* context

Assumptions (doubled)

- 2 independent SRS
- $n_1 p_1$
  $n_1(1-p_1)$ ≥ 10
  $n_2 p_2$
  $n_2(1-p_2)$
- $pop_1 \geq 10 \cdot n_1$
  $pop_2 \geq 10 \cdot n_2$

Example:

We are looking at the binge-drinking example again, except this time we are splitting the results by gender. We want to see if men and women are as likely to be binge drinkers. Use the results below to perform a full test of significance. $\alpha = 0.05$

	n	x	$\hat{p}$
Men	7180	1630	1630/7180 = 0.227
Women	9916	1684	1684/9916 = 0.1698
Total	17096	3314	

State

- 2 indep. SRS
- $n_1 p_1$, $n_1(1-p_1) \geq 10$, $n_2 p_2$, $n_2(1-p_2)$
- $pop_1 \geq 10 \cdot n_1$, $pop_2 \geq 10 \cdot n_2$

Check

- assumed
- 1630, 5550, 1684, 8232 ≥ 10 ✓
- $Pop_1 \geq 71800$, $Pop_2 \geq 99160$ ✓

X $H_0: p_M = p_F$

$H_a: p_M \neq p_F$

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = 9.337$$

$2 \cdot P(z > 9.337) = 1.011 \times 10^{-20}$

$2 \cdot P(z < -9.337) =$

We reject Ho b/c p-val < $\alpha = 0.05$. We have suff. evid. that prop. of male & female binge drinkers is not equal.

Find and interpret a 92% confidence interval for the difference in the proportion of college binge drinkers between men and women.

$$(\hat{p}_m - \hat{p}_F) \pm z^* \sqrt{\frac{\hat{p}_F(1-\hat{p}_F)}{n_F} + \frac{\hat{p}_M(1-\hat{p}_M)}{n_M}}$$

$$= (0.04631, 0.06808)$$

$(0.18, 0.20)$

We are 92% confident that the difference btw. the prop. of male & female binge drinkers is btw. 0.04631 and 0.06808.

Example #2:

Go back to your activity. Test to see if you quarter and penny have the same proportion of tails.

CALCULATOR:

Confidence Interval:

2 prop Z Int

Hypothesis Test:

2 prop Z test

#1 + 3

ASSUMP

(1) $$\left(\hat{p}_{Am} - \hat{p}_{pm}\right) \pm z^{*} \sqrt{\frac{\hat{p}_{Am}(1-\hat{p}_{Am})}{n_{Am}} + \frac{\hat{p}_{pM} - \hat{p}_{pm}}{n_{pm}}}$$

$$= (0.039, 0.201)$$

We are 90% confident that the diff between the prop. of day + night shift nurses that like their job is btw 0.039 + 0.201.

③ $H_0: p_s = p_c \leftarrow$ claim $p_s \le p_c$

Assump $H_a: p_s > p_c$

$$z = \frac{\hat{p}_s - \hat{p}_c}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = 2.037$$

$\hat{p} = \frac{x_1+x_2}{n_1+n_2}$ pooled estimator

$\alpha = 0.05$

$P(z > 2.037) = 0.0208$

We reject Ho in favor of Ha b/c p-value < α = 0.05.

We have suff. evidence that the prop. of defective items for the salesman is greater than the competitor's